A "Systematics" Tool for Medical Terminologies

Ying Tao, M.D., Eneida A. Mendonça, M.D., Ph.D., Yves A. Lussier, P. Eng., M.D.
Department of Medical Informatics, Columbia University, New York, NY, USA

Background. Finding the hierarchical relations amongst multiple terms within medical terminologies that support multiple parents to a term is a common task, especially for trainees and knowledge engineers implementing or maintaining medical logic modules or guidelines. Examples of such terminologies include the UMLS[®1] and the Medical Entity Dictionary (MED)[2]. In addition, the task of identifying and discriminating amongst some common ancestors to a list of terms is a recurrent theme. This is also a common concern in the science of classification (systematics). Some nearest common ancestors have distinct valuable properties for classification and simplification of lists. Although there exist some visualized navigating and editing tools for the UMLS and the MED, they are browsers that show a large number of unrelated and irrelevant relationships to the task at hand. While algorithms have been well studied in computer science to solve such a problem over semantic networks and trees[3], to our knowledge, they have not been used with a visualization tool in biomedicine.

We developed a visualized tool that graphically displays the hierarchical relations of multiple terms, and helps identifying the nearest common ancestors of these terms.

Methods. We used a Java Applet due to the requirement of graphic display and of intranet access. The UMLS and the MED were supported in the current version. The following steps are used by the system algorithm: a) do breath-first traversal according to child-parent relationship to generate a directed acyclic graph, b) do intersection search to find all the common ancestors of input nodes, c) find the nearest common ancestors (We do this by going through the marked common ancestors. If the term node doesn't have any children as common ancestors, we mark it as the nearest common ancestor), and d) mark common ancestors and nearest common ancestors using different colors on screen. We also integrated a MLM browser into this tool so that users can select the terminology codes in curly brackets directly by clicking on those codes. The tool takes plain text files as input and outputs the graph of relations in XML format.

Discussion. Figure 1 shows an example of an output generated by the tool. The example illustrates the identification of the common and the nearest common ancestors of 4 terms in the MED. Our
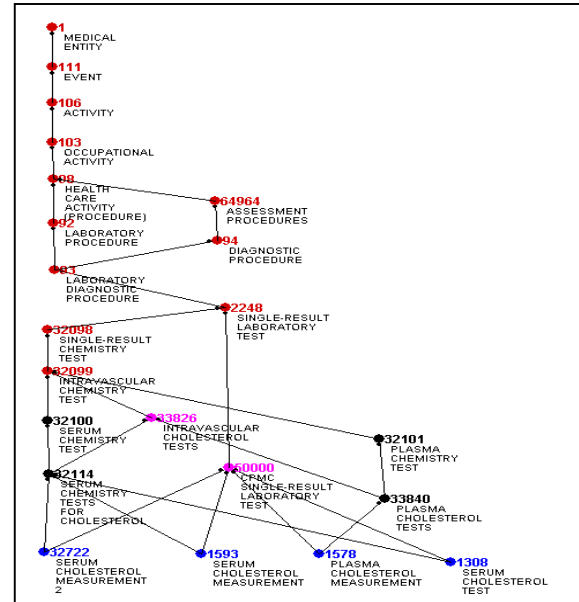

Figure 1. An output example showing the common ancestors (red) and the nearest common ancestors (pink) of four terms in the MED (blue). The black nodes correspond to other ancestors.

results demonstrate that the view shown is simpler than the alternative, a classical viewer of controlled terminologies. We believe that this tool can be very useful on tasks where the view of hierarchical relations of multiple terms is needed. There are, however, limitations. This method does not show other descendants of chosen common ancestors that are not part of the list of terms – an important finding to discriminate between equivalent choices in pink. Such algorithms are currently under investigation.

References

1. Bodenreider O. A semantic navigation tool for the UMLS. In: Proc AMIA Symp 2000; 2000. p. 971.
2. Cimino JJ. Terminology tools: state of the art and practical lessons. Methods Inf Med 2001;40(4):298-306.
3. Aho AV, Hopcroft JE, Ullman JD. On finding lowest common ancestor in trees. SIAM J. Comput. 1976;5(1):115-132.